**Lecture 22: Quaternary & Tertiary Structure Determination.**

Cox & Nelson: Section 3-3 and Box 4-4 (4e) or 4.3 (5e).



**Goal - Quaternary Structure Determination:**

i) Determine the subunit molecular weight.

ii) Determine the ratio of subunits.

iii) Determine the location of S-S bonds.

iv) Obtain overall native molecular weight.

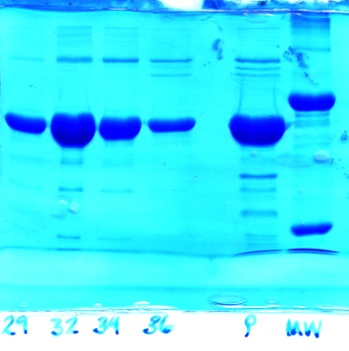
Propose a 4o structure consistent these data.

**SDS-polyacrylamide gel electrophoresis (PAGE).**

a) In the presence of an electric field proteins will migrate at a velocity that is proportional to their *intrinsic* charge-mass-ratio.

Top view of stained gel



b) **Denaturation** and coating of proteins by SDS gives them a *uniform* charge-to-mass ratio.



c) Forcing the protein-SDS complexes through a polyacrylamide gel causes separation of the proteins according to size.

d) Gels are stained with a stain that is specific for protein.

d) Distance, *d*, migrated down the gel during electrophoresis depends on molecular weight. **Smaller proteins move faster.**



**Denatured Molecular Weight (SDS-PAGE).**

i) Load molecular weight standards on one lane (right lane in above gel)

ii) load unknown in another lane.

iii) Electrophorese proteins through gel.

iv) Measure the distance migrated(d).

v) Plot log MW versus d for standards, generating a calibration curve.

vi) Use calibration curve to get logMW of unknown.

**Disulfide Bonds:** If proteins are crosslinked by disulfide bonds, and it is desirable to obtain the sizes of the subunits, then the S-S bonds have to be broken using β-mercaptoethanol (BME) or Dithiothreitol **(**DTT) before the electrophoretic separation. BME and DTT reduce the disulfide bond generating free –SH groups on the Cysteine residues.



**Native MW: Size exclusion (gel filtration).**

The beads in the gel-filtration column contains pores that allow smaller molecules to enter but exclude larger molecules. Gel filtration is usually performed under conditions where the quaternary structure of the protein is preserved, giving the **native molecular weight.**



The volume that a particular protein elutes from a column is called the elution volume, Ve. For example, if a protein was contained in the 69th mL of liquid that dripped from the column then its elution volume would be 69 mL.

A plot of log MW versus elution volume is linear. Proteins with known molecular weights are used to define the calibration curve, from which the logMW of the unknown can be obtained.

**Example:** An enzyme consists of four polypeptide chains. Two chains are 20 kDa in size (α-chain) and two are 30 kDa in size (β-chain). There is a single disulfide bond between the α and β subunits. The four chains associate as indicated in the diagram to form a hetero-tetramer, (αβ)2. The following two experiments were performed:



**1. Size exclusion chromatography**, using three known standards (IgG, Hb, Myo).

|  |  |  |
| --- | --- | --- |
| **Protein** | **MW (gm/mole (Da))** | **Log MW** |
| Antibody (IgG) | 150,000 | 5.17 |
| Unknown |  |  |
| Hemoglobin | 64,000 | 4.8 |
| Myoglobin | 16,000 | 4.2 |

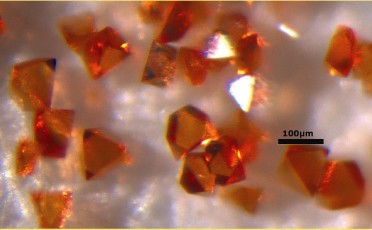
**2. SDS-PAGE** in the *absence* and *presence* of β-mercaptoethanol.

For calibration purposes, two proteins with known molecular weights, one with a molecular weight of 10 kDa and the other with a molecular weight of 160 kDa, were also included in this experiment. These two *standards* consist of a single polypeptide chain and will therefore give a single band in all experiments. Images of the two gels, as well as a plot of distance migrated versus log MW are shown.

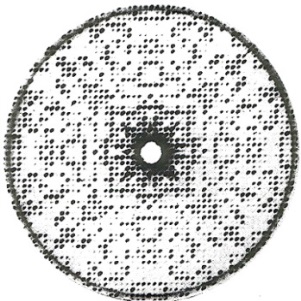
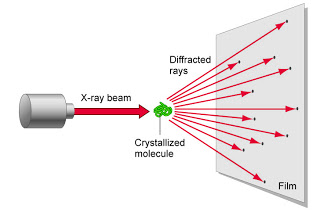
Note the 1:1 intensity of staining in the right gel = 1:1 α:β.

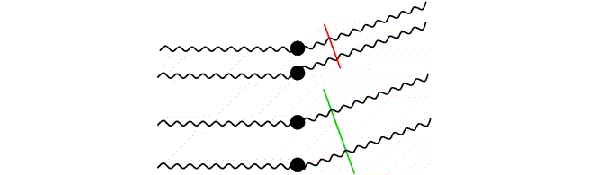


**Atomic Resolution Structures:**  **X-ray Diffraction**

1. Proteins must be crystallized in a regular lattice, just like a normal salt. The **unit cell** is the basic unit of a crystal. It is repeated over-and-over again to form the macroscopic crystal lattice. A single unit cell can contain multiple protein molecules.

(http://www.ruppweb.org/xray/tutorial/Crystal\_sym.htm)

2. X-rays are scattered by electrons – the amount of scattering is proportional to the number of electrons.

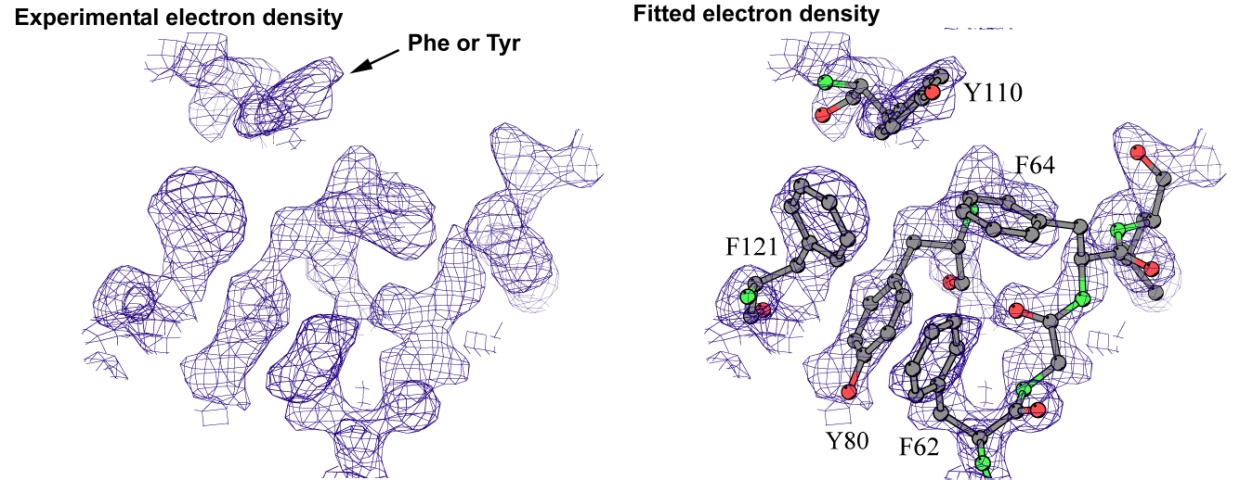
3. Interference between X-rays that are scattered from atoms in different locations changes the amplitude and the phase of the scattered X-rays. Therefore, scattered X-rays can be used to determine the position of atoms.

4. Intensities can be measured directly; phases have to be obtained indirectly. One common method of obtaining phases is called molecular replacement, where a homologous known structure is used to calculate the phases. Other methods include incorporating heavy atoms into the crystal to perturb the scattering.

5. A Fourier transform of the intensity and phases of the scattered X-rays produces an ‘electron density map’, or the number of electrons at each point in space in the crystal (ρ(x,y,z)). The crystallographer must figure out how to place, or "fit", the known primary structure of the protein into this map. This is aided by computational tools.

The overall process of obtaining X-ray structures is:

**Purify Protein→Grow Crystals→Collect Scattering Data→Determine Phase→Four. Trans→Fit Map.**



**Evaluation of structural quality (is the model correct?)**

* Ramachandran outliers – most (99%) of residues should fall in low energy region of plot.
* R-factor = (Predicted scattering – measured scattering)2. Should be 15-20%.
* Free R-factor = Use 90% of data to fit model. Calculate the R-factor with the 10% that was not used. This should be 20-25%.
* Free R – R factor ~ 5%, if higher, something **may** be wrong with the model.

*Why should you care?* Some structures are completely wrong (oops). Many have local errors.

**Accuracy of Atomic Positions:** Temperature factor (B) is a measure of mean atomic displacement: . B values > 75 Å2 = no electron density – no way to determine the structure in that region (1-10% of the residues, especially sidechains). B factors can be found in the file (PDB) that is being displayed by your software.

PDB File: ATOM 25 C GLY A 6 -4.478 67.908 10.486 1.00 20.15 C

*atom residue chain seq # x y z occ B*